心理科学进展 2021, Vol. 29, No. 3, 540–548
Advances in Psychological Science

https://dx.doi.org/10.3724/SP.J.1042.2021.00540

人们何以能够“舍生取义”？基于保护性价值观认知神经机制的解释*

岳 童[1] 黄希庭[1] 傅安国[2]

([1] 西南大学心理学与社会发展研究中心; 西南大学心理学部, 重庆 400715)
([2] 海南大学管理学院; 海南省公司治理研究院, 海口 570228)

摘 要 保护性价值观是一种拒绝与其他任何价值相互交易，尤其拒绝与经济价值进行交易的观念。近些年来的认知神经科学研究证据表明，人们之所以可以为了坚守保护性价值观，表现出“舍生取义”的行为(如对抗利益诱惑甚至放弃生命)，一方面在于保护性价值观是以一种道义主义的绝对规则方式进行表征和建构的，这使得人们较少进行利弊权衡；另一方面，由于保护性价值观与自我和道德认同过程紧密相关，使得其被主观赋予了最高价值。未来研究中，可以在借鉴保护性价值观的研究范式基础上，开展对中国人核心价值观的实证研究探讨，关注其文化差异问题，并进行正确价值观的干预和正确引导的应用实践研究。

关键词 保护性价值观，认知神经机制，道义主义，功利主义，逆火效应
分类号 B849:C91

1 引言

早在两千多年前，我国战国时期思想家孟子就曾提出过铿锵有力的价值判断：“生，亦我所欲也，义，亦我所欲也。二者不可得兼，舍生而取义者也” (《孟子·告子上》)。中国古代的儒家知识分子为了“行天下之大道”，要做到“富贵不能淫，贫贱不能移，威武不能屈”。这种对终极价值和信仰的坚守不仅护佑着中华民族成功地应对一场场生存危机，使无数仁人志士在面对利益诱惑甚至生命威胁时仍然可以保持初心，也是中华文明绵延至今的重要保障。无独有偶，西方宗教文化中为了信仰而放弃俗世生活(如“苦修主义”)，甚至“殉道”的例子也有很多，且这种对宗教信念的坚守长久以来被信徒们所推崇。以上做法不仅与经济学中的“理性人”假设相悖，而且不符合动物基本的趋利避害天性。为了某种价值信念而放弃现实利益甚至对抗死亡威胁，似乎是人类所特有的心理和行为模式，这体现了人类精神超越性的一面。

在心理学研究中，对上述现象的探讨往往在“保护性价值观(protected values)”的范畴内进行。保护性价值观也被称为神圣价值观(sacred values)或绝对性价值观(absolute values)，是具有以下性质的价值观念：人们拒绝将这种观念指向的客体与具有价值的其他客体互相交易，尤其拒绝与那些仅仅具有经济价值的客体相交易(Baron & Spranca, 1997; Baron & Jonathan, 2017; 何贵兵, 奚岩, 2005)。在实际操作中，研究者主要通过两种方式来界定某种价值观是否属于保护性价值观：第一种是通过问卷测量的方式(Dogan et al., 2016; Duc et al., 2013)，主要采用的工具是保护性价值观问卷(Hanselmann & Tanner, 2008; Tanner et al., 2009)，它包涵9个题项，其中4个直接测量个体在某种观念上抵制金钱交易的倾向，另外5个题项所衡量的是个体对某种价值观被违反之后的情绪反对倾向，最后在该量表上的总得分便是某种价值观的保护性程度；第二种名为Becker-DeGroot-Marshak (BDM)拍卖任务，即让被试对自己所有的价值观进行金钱标价，对那些不论给

收稿日期：2020-06-02
* 西南大学博士启动基金项目(swu118092)、重庆市社科规划办社科基金项目(2018PY81)和国家社科基金青年项目(19CSH051)资助。
通信作者：黄希庭, E-mail: xthuang@swu.edu.cn

出多高的价格被试都不会出卖的价值观便被定义为保护性价值观(Berns, Bell, et al., 2012; Pincus et al., 2014)。从以上操作性定义来看, 保护性价值观包括基本宗教信念、道德规范[如诚信(Dogan et al., 2016)、正直(Berns, Bell, et al., 2012)]以及有关国家与种族认同的核心观念等(Berns, Prietula, et al., 2012; 李林, 黄希庭, 2013)。需要指出的是, 价值观并非是完全普适性的, 而是会受到个人所处文化环境和成长经历的影响。例如, 以国外大学生为被试的研究中发现, 神学专业的学生相比于经济学专业的学生有更多的保护性价值观(Duc et al., 2013; Tanner et al., 2009)。总之, 保护性价值观可以视为个体某些价值观在特定经历和文化塑造下的绝对化和神圣化过程, 可以为人们终极人生目标的达成提供核心和持久的动力。

自保护性价值观这一概念进入心理学研究领域以来, 围绕着该心理现象有何加工特点和社会功能问题, 研究者进行了大量的探讨。例如, 很多研究者发现当用金钱补偿来迫使个体在某些保护性价值观上进行妥协时, 个体不仅会产生一系列负性情绪(如愤怒、厌恶等) (Atran & Ginges, 2012; Sheikh et al., 2012)来对抗这种交易, 还会进一步加强对自己原有价值观的坚持, 即形成所谓的“逆火效应” (Backfire Effect) (Bénabou & Tirole, 2011; Sandel, 2012)。不仅如此, 由于不同群体所处的历史文化背景不同, 其所持有的保护性价值观存在很大的差异, 而这些价值观又具有绝对性和现实超越性, 进而导致不同群体之间的理解和对话变得非常困难, 这也成为了文化和群际冲突频发的重要原因之一(Ginges et al., 2007; Atran & Ginges, 2012 ;Sheikh et al., 2016)。然而, 上述研究很大程度上都是从现象层面对保护性价值观进行描述和总结, 究竟保护性价值观是如何对抗利弊权衡(如金钱诱惑或生命威胁), 从而表现出绝对性和现实超越性的, 以往的研究者很难给出令人满意的回答。近些年来, 随着认知神经科学的兴起及功能性磁共振成像(functional magnetic resonance imaging, fMRI)技术的发展, 揭示保护性价值观的心理及神经加工机制成为了可能(Vilarroya & Hilferty, 2013), 这也为理解为何人们能够“舍生取义” (即放弃趋利避害的本性而做出坚守某种价值理念的行为)提供了更为直观的分析视角和证据。有鉴于此, 本文将在回顾和分析前人研究基础上, 首先论述保护性价值观的表征和建构模式, 之后从认知神经加工的角度阐释保护性价值观对抗利弊权衡的作用机制问题, 最后进行总结并提出未来研究的展望, 以期可以更为深入系统地理解保护性价值观的作用原理, 并为应用实践中正确价值观的培养和干预提供理论上的启示。

2　非保护性和保护性价值观*的表征及建构模式

价值观从本质上来说是一系列抽象的观念表征(黄希庭, 2014)和社会规则的集合(Zahn et al., 2007)。而从已有研究来看, 人们基于价值观念的决策判断主要分为两种：功利主义(utilitarian)和道义主义(deontological)。前者是在效用框架下进行的, 往往在利弊权衡后按照利益最大化的原则做出决策; 而后者更注重对事件性质而非结果的判断, 且只有绝对正确或错误的二元性区分, 与道德直觉(moral intuition)相似(Baron, 2009)。其中, 很多非保护性价值观是基于功利主义原则进行加工的, 而保护性价值观则更大程度上是以一种道义主义的方式进行表征和建构的。

2.1　非保护性价值观：利弊权衡后的价值最大化加工

在很多情况下, 价值观并非是绝对稳定存在的心理结构。有研究者认为, 价值观是自明之理, 缺乏必要认知成分的支持, 这使得某些非保护性价值观的重要性容易受到外界劝说和社会压力的影响(Hitlin & Piliavin, 2004)。而最近的脑成像研究结果发现, 个体之所以能在外界环境影响之下改变自己的某些非保护性价值观, 其背后涉及到大脑价值赋值系统的参与。

腹内侧前额叶(ventromedial prefrontal cortex, vmPFC)和腹侧纹状体(ventral striatum, VS)是人类价值赋值的两个核心脑区, 在加工对个体来说

* 需要指出的是, “非保护性价值观”并非一个专有的学术名词, 而是在本文中为了与保护性价值观相区分而人为划定的概念。从包涵的内容来看, 保护性价值观仅仅在价值观系统中占很少一部分, 而大多数价值观属于非保护性的。而也并非所有的非保护性价值观都会采用功利主义的加工方式, 从功利主义到道义主义更像是一个连续体的两级。因此, 在论述非保护性价值观的加工模式时, 本文采用了“某些非保护性价值观”的说法。

有奖赏性质[包括初级(如食物)和次级(如金钱)]的刺激中发挥着非常重要的作用(Bartra et al., 2013; Levy & Glimcher, 2012; McNamee Rangel & O'Doherty, 2013)。而人们之所以在他人劝说或社会压力下改变自己原本持有的某些非保护性价值观，很大程度上是出于利益最大化的考虑，即放弃固有价值观并遵从被主观赋予了更大经济或生存价值的新选择。例如，在 Cooper 等人(2017)的研究中发现，当劝说久坐人群改变其不健康的生活方式并接受健康生活理念时，他们价值计算系统的功能连接强度与在实验后久坐行为的频率减少有关：腹内侧前额叶和腹侧纹状体在“增加身体活动”的健康观念的刺激下激活及连接水平越强，随后久坐时间缩减幅度越大。这说明，健康生活的观念可能会使某些久坐个体意识到健康的好处，在大脑中进行权衡，进而根据自我利益最大化的原则更改自己原有的生活理念。再如，人类是社会性动物，与群体内大部分成员保持观点一致在多数情况下都是最优化的选择，这对个体生存至关重要。Nook 和 Zaki (2015)考察了个体对事物的健康理念是如何被同伴的选择所影响的，结果发现当被试提升自己对食物的健康评价以符合同伴观念的过程中，其腹内侧前额叶的活动也随之增强。该研究结果表明，个体在同伴选择压力下会进行价值权衡和计算，并做出观念上的改变以使自己获得现实的利益。

由上可知，虽然价值观是人们在社会化过程中所形成的理念和规则，能在默认模式下直接影响生活中的种种决策。但是在面对利弊权衡时，很多原本持有的价值规则也要让位于对现实利益的考量，即趋利避害在很多情况下是个体的最优价值选择。然而，如果在所有价值选择上个体都遵循上述加工模式的话，也就没有“杀身成仁，舍生取义”这种超越性价值信念存在的可能了。也就是说，保护性价值观能够抵御或越过利弊权衡过程，不计自身得失而坚持相应的信念，在其表征及建构模式上必然有其独特之处。

2.2 保护性价值观：基于道义原则的绝对规则加工

自保护性价值观的概念提出之日起，相关研究者便认为，这种价值观与非保护性价值观有所不同，即它们可能更多是以道义取向进行表征和建构的，不问结果如何，只从事件本身的对与错出发进行价值判断。近些年来，这种观点在神经生理层面得到越来越多证据的支持。

例如，Berns, Bell 等人(2012)发现，被试对保护性价值观的表征和处理主要激活的脑区为腹外侧前额叶(ventrolateral prefrontal cortex, vlPFC)和颞顶联合区(temporo-parietal junction, TPJ)，其中前者与语义规则的提取与加工有关(Sharot et al., 2012)，而后者主要参与到了对信念的道德判断过程(Young et al., 2010)。与之不同的是，人们对非保护性价值观的加工则更多地利用顶下小叶(inferior parietal lobules, IPL)这一与功利主义的价值计算与决策有关的脑区(Bhanji et al., 2010; Kable & Glimcher, 2009)。随后，Pincus 等人(2014)也发现了类似的结论，即被试在加工涉及保护性价值观的陈述时相比于非保护性价值观更多地激活了腹外侧前额叶这一与规则提取相关的脑区，这再次说明保护性价值观是按照一种绝对规则的方式存储于大脑之中的。不仅如此，Pincus 等人(2014)还试图通过同辈的选择一致性压力(指周围同伴都不认可自己选择时所带来的心理压力)来影响个体对价值观的坚持，结果发现越不容易受到社会压力影响来改变固有价值理念的被试，其腹外侧前额叶的激活程度越强，这可能表明不同个体在保护性价值观的规则认同程度上也存在差异：越是将其当成一种道义原则进行表征和构建的个体，保护性价值观于他们来说越是稳固。总之，以上结果表明，保护性价值观在表征和建构模式上的确有其独特之处，这种认知神经结构决定了人们在加工保护性价值观时采用了一种绝对规则的方式，而非权衡利弊再进行判断。

另外，保护性价值观的表征和建构除了与道义性的绝对规则加工方式有关之外，可能还涉及到自我表征和情绪加工过程。例如，Kaplan, Gimbel 和 Dehghani 等人(2016)从超过 2000 万的网络日志中筛选出包涵保护性和非保护性价值观的真实生活故事呈现给被试，并记录了他们在这个过程中的神经反应。结果发现，当个体阅读这些故事时，其后内侧皮质(posterior medial cortices, PMC)，内侧前额叶皮层(medial prefrontal cortex, mPFC)及颞顶联合区等默认网络相关的脑区有明显的激活，且包含保护性价值观的故事相比于包含非保护性价值观的故事在上述脑区上激活程度更甚。Kaplan, Gimbel 和 Dehghani 等人认为，上述结果说明了保护性价值观的表征可能涉及到大

脑的自我身份认同、道德判断和社会情绪加工过程。随后, Kaplan, Gimbel 和 Harris (2016)考察了当个体强烈认同的保护性价值观(自由主义 vs 保守主义)受到反对信息的冲击时, 大脑是如何进行处理和加工的。结果也发现, 相比于非保护性的价值观, 被试在加工与政治信念相关的信息时其默认网络相关脑区有明显的激活。从个体差异的角度上来看, Kaplan, Gimbel 和 Harris 还发现了越是能在反对信息冲击下坚持原本价值信念的个体, 其背内侧前额叶皮层(dorsomedial prefrontal cortex, dMPFC)、脑岛(insula)和杏仁核(amygdala)的活动程度越强。对此, Kaplan, Gimbel 和 Harris 认为这可能反映了保护性信念在维持过程中个体是如何利用情绪唤起来对抗威胁信息冲击的。综上所述, 保护性价值观在表征和建构上可能还与个体的自我与道德认同有关, 这就使得它们在价值重要性上明显提升, 且可以调动情绪资源来对抗外界信息的冲击。

3 “舍生取义”: 保护性价值观对抗利弊权衡的认知神经机制

对个体来说意义最为重大的需求往往来源于两个方面: 生存需要和物质利益。前者保障了个体生命的延续, 后者可以使个体获得良好的发展。如果是基于利益最大化的原则, 避免死亡威胁和获得现实物质奖赏是生物体在趋利避害本能驱使下的不二选择。然而, 保护性价值观却可以使人们为了某种信念抵御物质利益的诱惑, 甚至不惜放弃自己的生命, 即人们所说的“舍生取义”。如前文所述, 保护性价值观具有独特的表征和建构模式来使其保持稳定性, 但这种心理结构是如何对抗物质诱惑和死亡威胁的呢? 近些年来的认知神经研究尝试从以上两个方面进行了探讨, 这也为人们理解保护性价值观的工作原理和内在机制提供了新的视角。

3.1 保护性价值观如何对抗物质诱惑

保护性价值观的一个最典型特点便是不可交易性, 即不论提出多少物质补偿个体都不愿放弃自己坚持的价值理念。之所以如此, 一方面在于保护性价值观基于绝对规则的构建方式, 另一方面则可能源于价值观被冲击后产生的情绪对抗过程, 即前文所提及的“逆火效应”。

Dogan 等人(2016)以“诚信”作为保护性价值观的代表, 探讨了个体是如何抵御金钱诱惑的。他们要求被试扮演 CEO 的角色, 并在诚实报告公司收入或虚报公司收入以获得个人财富提升之间进行选择。脑成像研究结果表明, 被试越是将诚信视为自己的保护性价值观, 在面对金钱诱惑时其背外侧前额叶、背内侧前额叶及其额下回(inferior frontalis gyrus, IFG)的激活水平越强。进一步分析结果表明, 额下回与背外侧前额叶(dorsolateral prefrontal cortex, dlPFC)的单向连接及其额下回与背内侧前额叶的双向连接越强, 被试将诚信价值观与金钱收益交换的意图越弱。其中, 背外侧前额叶、背内侧前额叶及其额下回在之前的研究中都参与到了认知控制及其反应抑制过程(Aron, 2007; Carter & Veen, 2007), 说明那些更加重视诚信价值观的个体应当具备更强的认知控制能力来抑制金钱的诱惑; 而额下回在之前的研究中多与语义规则的表征有关(Souza et al., 2009), 这可能说明以诚信为保护性价值观的被试相关价值信念在其语义规则网络中更加稳固。上述研究可能表明, 对价值观规则性的绝对认知及其主动抑制可能是个体能在金钱诱惑下坚持保护性价值观的重要原因。

Duc 等人(2013)探讨了保护性价值观强制交易引发的“逆火效应”及其背后的神经反应过程。他们召集了经济学和神学专业的被试来完成保护性价值观交易任务, 如在人的生命和金钱补偿中进行强制性选择。其中, 神学专业的学生更倾向于将人的生命权视为保护性价值观, 因此他们在面对交易时报告出了更高水平的愤怒与厌恶情绪, 在神经活动上更多地激活了右侧杏仁核(amygdala)和前颞叶(anterior temporal lobe, aTL)等脑区, 且其内侧前额叶−颞叶−边缘系统网络及杏仁核之间的连接强度与其自我报告的负性情绪水平之间存在更强的相关。参考前人研究结果, 杏仁核及其额叶−颞叶−边缘系统网络更强的活动可能说明神学专业学生在交易选择后会产生更强的情绪关注水平, 而前颞叶等脑区的激活可能反映了道德相关的规则被违反时复杂的知觉过程(Zahn et al., 2007)。总之, 上述结果说明, 价值规则被违反后所激发的负性情绪体验可能是保护性价值观对抗金钱诱惑、维持自身合理性的重要途径。

综上所述, 虽然目前直接探讨保护性价值观如何对抗物质诱惑的神经生理研究并不多, 但结

合前文所论述的保护性价值观的表征和构建模式，可以大致推测和勾勒出其可能的运行机制。首先，保护性价值观是以一种基于道义主义的绝对规则形式来构建的，这就在很大程度上保证了这种观念在面对物质诱惑时可以绕开利弊计算过程，直接从语义记忆网络中提取相关规则来遵守和进行判断；另外，由于保护性价值观在表征和构建上与自我与道德认同有关，因此在强制要求个体在价值信念和金钱利益之间进行选择时，可能会被个体视为对自我和道德认同的威胁，因此激发起愤怒和厌恶等消极情绪反应进行反抗，进而巩固自己原本持有的信念。

3.2 保护性价值观如何超越求生本能

对于生物体而言，避免死亡的决策往往是最优价值选择。然而，很多人可以为了坚持自己的价值理念，宁可放弃生命权，这种现象在宗教极端团队中较为常见。最近，很多研究从大脑神经活动的角度揭示了宗教极端团体的成员或支持者为何可以为了保护性价值观牺牲生命的心理过程，从中也可以给我们一定的启发。

已有研究结果发现，宗教极端团体成员在面对保护性价值观和为之牺牲生命的价值选择时，绝对性的规则加工方式仍然在起主要作用，而利弊计算则较少被使用。例如，Pretus 等人(2018)对巴塞罗那年轻的摩洛哥裔男子进行研究，这些被试都公开表示他们将不惜以生命来捍卫圣战事业，即宗教教义对他们来说是保护性的价值观。在进行大脑扫描时，Pretus 等人(2018)测量了这些被试为自己的保护性价值观(例如，禁止关于先知的漫画，禁止同性恋婚姻)及其他重要但非保护性的价值观(如穿着面纱的女性，在学校中的伊斯兰教学等)而战斗或赴死的意愿。行为结果发现，相比于非保护性的价值观，被试对保护性价值观报告出了更强的战斗和献身意愿。在大脑活动水平上，保护性价值观相比于非保护性价值观更多地激活了额下回(IFG)这一与语义规则提取和加工相关的脑区，这说明保护性价值观在这些被试认知结构中是一种绝对性的规则，在面对死亡交换时仍然被当成不可违背的规则来进行坚持。Hamid 等人(2019)通过访谈招募了在西班牙的巴基斯坦移民中的宗教狂热分子，这些人都是虔诚军(Lashkar-et-Taiba)的拥护者，因此在与宗教相关的价值观上非常极端。在实验中，给上述被试呈现保护性和非保护性的价值观，并要求他们报告自己为该理念战斗或赴死的意愿，结果发现保护性价值观相比于非保护性价值观在背外侧前额皮层上激活更弱。背外侧前额皮层涉及到执行控制相关的心理功能，如外在刺激的利弊评估(Rudorf & Hare, 2014)和损益计算(Hare, Camerer & Rangel, 2009)。由此 Hamid 等人(2019)认为，当人们选择为保护性价值观献出生命时，他们更少地进行利弊计算，而是更多地将保护性价值观视为引导目标的道德要求。

Pretus 等人(2019)的研究结果为该问题提供了更进一步的启发。他们在巴塞罗那招募了巴基斯坦的穆斯林移民参加实验，并在其中筛选出圣战组织和民族主义的坚定支持者作为具有宗教极端保护性价值观的典型代表。在实验中，给这些被试呈现“西方军队应该被驱逐出穆斯林国家”、“印度不能拥有克什米尔的主权”等涉及保护性价值观的陈述，然后要求这些被试在 7 点 Likert 量表上进行“是否愿意为该理念战斗或献身”的等级评定。结果发现，被试愿意为该价值理念战斗和赴死的意愿越高，其腹内侧前额叶的活动水平越强，背外侧前额皮层的激活程度越弱，且这两个脑区的功能连接倾向也越不明显。如前文所述，腹内侧前额叶与决策过程里的主观价值赋值有关(Bartra et al., 2013; Shenhav & Greene, 2014)，而背外侧前额皮层与协商推理及整合成本效益计算相关。这可能说明，这些极端主义分子将为信念而牺牲生命主观赋予了最高的价值，利弊计算和认知控制通路在其中基本是不起作用的。作为对照，Pretus 等人(2019)在研究的第二部分故意设置让这些被试看到身边同伴的低战斗和赴死意愿，以此来观测上述两个脑区的活动模式。结果发现，被试受到了同伴观点的影响，在为信念放弃生命的主观报告意愿上有了明显降低，伴随这种趋势的是背外侧前额皮层的活动增加。以上研究结果表明，对保护性价值观的主观赋值和利弊计算之间可能存在着一种此消彼长的拮抗关系。

总之，从已有结果来看，某些与宗教信仰有关的保护性价值观之所以能够超越求生本能，可能还是与其独特的表征和建构模式密不可分。一方面，作为绝对性的道义规则，即使在面对为信念牺牲的可能性时保护性价值观也是人们的第一选择，而不需要进行利弊权衡；另一方面，保护

性价值观与自我和道德认同在表征上的一致性使其成为了个体生活中的最高价值目标，这使得利弊计算系统很难在这个过程中发挥作用。

4　总结与展望

在人类历史发展长河中，不论在何种文化背景下，为了坚持自己理想信念而"富贵不能淫，贫贱不能移，威武不能屈"的行为都广受推崇和歌颂。这种行为使人们超越了趋利避害的生物性本能，是人类心理独特性的所在。近些年来关于保护性价值观认知神经机制的研究，为人们理解"舍生取义"背后的心理加工过程提供了神经生理层面的证据。从已有研究结果来看，人们之所以可以为了某种坚守的信念而对抗物质诱惑甚至舍弃生命，一方面在于保护性价值观是以一种道义主义的绝对规则方式进行表征和建构的，这使得人们在进行比较和决策时更多地只需要遵守，而较少进行利弊权衡；另一方面，由于保护性价值观与自我和道德认同过程紧密相关，这使得保护性价值观被主观赋予了最高价值，且一旦被强制纳入到利弊计算模式还会激发个体反抗性的愤怒情绪，从而进一步加强了人们对这种信念的坚守。总的来说，目前对该问题的探讨正处于刚刚起步的阶段，尤其国内对该领域的实证研究还比较缺乏，未来可能还需要在以下几个方面做进一步探讨。

首先，可以借鉴保护性价值观的研究范式和思路，开展对中国人核心价值观的实证研究探讨。当前，如何培育和践行"诚信"、"爱国"等社会主义核心价值观，使之能够获得全社会的广泛认同并能在教育、精神文明创建及社会发展的方方面面中发挥引领作用，是广大学者和教育工作者长期关注的问题。而这种认同和转化的关键，便是如何将社会主义核心价值观纳入到个体的核心价值观中。然而，对于中国人核心价值观的结构、作用机制甚至神经表征问题，目前都很少有研究者进行过探讨。虽然保护性价值观和中国人所说的核心价值观是不同的概念，但可以借鉴国外对保护性价值观认知神经机制探讨的范式和思路，开展对中国人核心价值观的实证研究。具体而言，可以从动态加工过程和静态结构两个方面进行推进。一方面，基于前文论述可知，保护性价值观得以对抗利弊权衡，可能是"冷"认知规则系统和"热"主观赋值和情感激发系统的相互作用的结果：前者给个体提供最直接的价值选择，后者是通过自我认同、道德情感等方式帮助个体忽视现实利益的考虑。核心价值观也有类似的功能，那么其作用模式是否也与保护性价值观相似？未来研究中可以参考相关实验思路进行检验。另一方面，前人研究结果已经发现价值观具有稳定的神经基础，在大脑功能的自发活动甚至结构层面都有所体现(Zacharopoulos et al., 2016; Zacharopoulos et al., 2017)。鉴于核心价值观在人格系统中的稳定地位，未来研究可以借鉴相关技术，如探索其在大脑静息态和结构态上的存在和表现形式，以达到对这一问题更为深入的认识。

其次，在借鉴国外保护性价值观的研究思路时，应充分意识到文化差异和内涵所指不同的问题，不能盲目照搬。目前，对保护性价值观的研究主要是在西方文化背景下进行的。而 Kaplan, Gimbel 和 Dehghani 等人的研究发现，伊朗、美国和中国被试在加工保护性价值观过程中大脑激活模式存在显著差异，其中中国被试在加工保护性价值观时与非保护性价值观脑激活差异最小，他们认为这可能是因为在东亚文化下的被试更强调辩证思维模式和背景依存型的自我建构方式。但实际上，上述结论很可能是因为文化背景不同所导致的：保护性价值观是从宗教中的"神圣价值观"中发展而来的，在这种内涵界定下，普遍持无神论的中国被试很难像伊朗和美国被试一样表现出明显的保护性价值观加工模式。这提醒我们，西方文化下所界定的保护性价值观和中国人所认为的应当受保护的价值观存在明显的不同。以本文所举的"舍生取义"为例，中国传统文化下的"义"当表述为超越自我、面向现实世界的正义，是人间正道的准则，具有强烈的道德涵义(荆惠民, 2006)，是不可能包括类似宗教狂热分子所持的极端价值观念的。因此，在国外研究框架内所探讨的保护性价值观(尤其是与宗教有关的保护性价值观)和中国文化中所提倡的能够使无数仁人志士安身立命的保护性价值观，不论是在内涵还是作用机理上都有显著不同。这样，完全在西方保护性价值观的语境和框架内进行研究，盲目照搬而不考虑中国实际情况便存在极大的问题。所以，今后在中国文化和社会环境下探讨保护性价值观问题时，需要深入思考其内涵究竟为何，防止削

足适履，以西方文化下的概念来强行解释中国人的心理。

最后，需要科学开展对中国人保护性价值观或核心价值观的干预和正确引导的应用实践研究。价值观干预和引导的第一步，在于界定何为正确的价值观。然而，在当前的操作性定义之下(拒绝与那些仅仅具有经济价值的客体相交易的价值观)，保护性价值观是一个中性的概念，因此宗教极端团体所持有的狂热理念也可以被囊括入内。而在中国学者对价值观的主流定义中，价值观是有正确与错误之区别的，是人们用来区分好坏、美丑、益损、正确与错误的信念系统(黄希庭, 2014)。这里便存在着价值观研究和界定中“形式”和“内容”之间关系的问题：仅仅是简单粗暴地用操作性定义划定一条是或否的区分线，还是应当赋予其更多的价值和社会引导内涵。无疑，当前研究中对保护性价值观的操作性定义使得这一问题更容易在实验室中进行经验验证(舒跃育 等, 2019)，但将价值观还原为市场价值隐喻，会使人本身失去了其内在固有价值。因此，今后在借鉴国外保护性价值观的研究成果，进而对正确核心价值观的认同和培养开展干预和引导的应用实践过程中，需要明确价值观界定中“形式”和“内容”之间的关系，不能用已有的操作性定义一刀切，而是需要更为深入地思考和更加科学精准的定义，并探寻其背后的形成作用机理，如此才能提升价值观在行为塑造和社会心态引领中的积极作用。

参考文献

何贵兵, 奚岩. (2005). 保护性价值观及其对决策行为的影响. *应用心理学, 11*(1), 60–66.

黄希庭. (2014). *探究人格奥秘*. 北京：商务印书馆.

荆惠民. (主编). (2006). *中国人的美德：仁义礼智信*. 北京：中国人民大学出版社.

李林, 黄希庭. (2013). 价值观的神经机制：另一种研究视角. *心理科学进展, 21*(8), 1400–1407.

舒跃育, 石莹波, 袁彦. (2019). “操作性定义”和“证伪标准”不足以为心理学奠基. *心理学报, 51*(9), 1068–1078.

Aron, A. R. (2007). The neural basis of inhibition in cognitive control. *The Neuroscientist, 13*(3), 214–228.

Atran, S., & Ginges, J. (2012). Religious and sacred imperatives in human conflict. *Science, 336*(6083), 855–857.

Baron, J. (2009). Belief overkill in political judgments. *Ssrn Electronic Journal, 29*(4), 368–378.

Baron, J., & Spranca, M. (1997). Protected values. *Organizational Behavior & Human Decision Processes, 70*(1), 1–16.

Baron, J., & Jonathan. (2017). Protected values and other types of values. *Analyse & Kritik, 39*(1), 85–100.

Bartra, O., Mcguire, J. T., & Kable, J. W. (2013). The valuation system: A coordinate-based metaanalysis of bold fmri experiments examining the neural correlates of subjective value. *NeuroImage, 76*, 412–427.

Bénabou, R., & Tirole, J. (2011). Identity, morals, and taboos: Beliefs as assets. *The Quarterly Journal of Economics, 126*(2), 805–855.

Berns, G. S., Bell, E., Capra, C. M., Prietula, M. J., Moore, S., Anderson, B., … Atran, S. (2012). The price of your soul: Neural evidence for the non-utilitarian representation of sacred values. *Philosophical Transactions of the Royal Society B: Biological Sciences, 367*, 754–762.

Berns, G. S., Prietula, M., & Lemieux, A. F. (2012). Neurobiology of sacred values and implications for radicalization process. *National Security Challenges: Insights from Social, Neurobiological, and Complexity Sciences, 28*, 164–172.

Bhanji, J. P., Beer, J. S., & Bunge, S. A. (2010). Taking a gamble or playing by the rules: Dissociable prefrontal systems implicated in probabilistic versus deterministic rule-based decisions. *NeuroImage, 49*(2), 1810–1819.

Carter, C. S., & van Veen, V. (2007). Anterior cingulate cortex and conflict detection: An update of theory and data. *Cognitive, Affective, & Behavioral Neuroscience, 7*(4), 367–379.

Cooper, N., Bassett, D. S., & Falk, E. B. (2017). Coherent activity between brain regions that code for value is linked to the malleability of human behavior. *Scientific Reports, 7*, 43250.

Dogan, A., Morishima, Y., Heise, F., Tanner, C., Gibson, R., Wagner, A. F., & Tobler, P. N. (2016). Prefrontal connections express individual differences in intrinsic resistance to trading off honesty values against economic benefits. *Scientific Reports, 6*(1), 33263–33263.

Duc, C., Hanselmann, M., Boesiger, P., & Tanner, C. (2013). Sacred values: Trade-off type matters. *Journal of Neuroscience, Psychology, and Economics, 6*(4), 252–263.

Ginges, J., Atran, S., Medin, D., & Shikaki, K. (2007). Sacred bounds on rational resolution of violent political conflict. *Proceedings of the National Academy of Sciences, 104*, 7357–7360

Hamid, N., Pretus, C., Atran, S., Crockett, M. J., Ginges, J., Sheikh, H., … Vilarroya, O. (2019). Neuroimaging ‘will to fight’ for sacred values: An empirical case study with supporters of an Al Qaeda associate. *Royal Society Open Science, 6*(6).

Hanselmann, M., & Tanner, C. (2008). Taboos and conflicts in decision making: Sacred values, decision difficulty, and emotions. *Judgment and Decision Making, 3*, 51–63.

Hare, T. A., Camerer, C. F., & Rangel, A. (2009). Self-control in decision-making involves modulation of the vmPFC valuation system. *Science, 324*(5927), 646–648.

Hitlin, S., & Piliavin, J. A. (2004). Values: Reviving a dormant concept. *Annual Review of Sociology, 30*(1), 359–393.

Kable, J. W., & Glimcher, P. W. (2009). The neurobiology of decision: Consensus and controversy. *Neuron, 63*(6), 733–745.

Kaplan, J. T., Gimbel, S. I., Dehghani, M., Immordinoyang, M. H., Sagae, K., Wong, J. D., … Damasio, A. R. (2016a). Processing narratives concerning protected values: A cross-cultural investigation of neural correlates. *Cerebral Cortex, 27*(2), 1428–1438.

Kaplan, J. T., Gimbel, S. I., & Harris, S. (2016b). Neural correlates of maintaining one's political beliefs in the face of counterevidence. *Scientific Reports, 6*(1).

Levy, D. J., & Glimcher, P. W. (2012). The root of all value: A neural common currency for choice. *Current Opinion in Neurobiology, 22*(6), 1027–1038

McNamee, D., Rangel, A., & O'Doherty, J. P. (2013). Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. *Nature Neuroscience, 16*(4), 479–485

Nook, E. C., & Zaki, J. (2015). Social norms shift behavioral and neural responses to foods. *Journal of Cognitive Neuroscience, 27*(7), 1412–1426.

Pincus, M., Laviers, L., Prietula, M. J., & Berns, G. S. (2014). The conforming brain and deontological resolve. *Plos One, 9*(8), e106061.

Pretus, C., Hamid, N., Sheikh, H., Ginges, J., Tobeña, A., Vilarroya, O., … Atran, S. (2018). Neural and behavioral correlates of sacred values and vulnerability to violent extremism. *Frontiers in Psychology, 9*: 2462.

Pretus, C., Hamid, N., Sheikh, H., Gomez, A., Ginges, J., Tobena, A., … Atran, S. (2019). Ventromedial and dorsolateral prefrontal interactions underlie will to fight and die for a cause. *Social Cognitive and Affective Neuroscience, 14*(6), 569–577.

Rudorf, S., & Hare, T. A. (2014). Interactions between dorsolateral and ventromedial prefrontal cortex underlie context-dependent stimulus valuation in goal-directed choice. *The Journal of Neuroscience, 34*(48), 15988–15996.

Sandel, M. (2012). *What Money Can't Buy: the Moral Limits of Markets*. New York: Farrar, Straus and Giroux.

Sharot, T., Kanai, R., Marston, D., Korn, C. W., Rees, G., & Dolan, R. J. (2012). Selectively altering belief formation in the human brain. *Proceedings of the National Academy of Sciences, 109*(42), 17058–17062.

Sheikh, H., Ginges, J., Coman, A., & Atran, S. (2012). Religion, group threat and sacred values. *Judgment and Decision Making, 7*(2), 110–118.

Sheikh, H., Gómez, Á., & Atran, S. (2016). Empirical evidence for the devoted actor model. *Current Anthropology, 57*, S204–209.

Shenhav, A., Greene, J. D. (2014). Integrative moral judgment: Dissociating the roles of the amygdala and ventromedial prefrontal cortex. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience, 34*, 4741–4749.

Souza, M. J., Donohue, S. E., & Bunge, S. A. (2009). Controlled retrieval and selection of action-relevant knowledge mediated by partially overlapping regions in left ventrolateral prefrontal cortex. *Neuroimage, 46*(1), 299–307.

Tanner, C. (2009). To act or not to act: Nonconsequentialism in environmental decision-making. *Ethics & Behavior, 19*, 479–495.

Vilarroya, O., & Hilferty, J. (2013). The neuroimaging of sacred values. *Annals of the New York Academy of Sciences, 1299*, 25–35.

Zacharopoulos, G., Hanel, P. H., Lancaster, T. M., Ihssen, N., Drakesmith, M., Foley, S., … Linden, D. E. (2017). Nonlinear associations between human values and neuroanatomy. *Social Neuroscience, 12*(6), 673–684.

Zacharopoulos, G., Lancaster, T. M., Bracht, T., Ihssen, N., Maio, G. R., & Linden, D. E. J. (2016). A hedonism hub in the human brain. *Cerebral Cortex, 26*(10), 3921–3927.

Zahn, R., Moll, J., Krueger, F., Huey, E. D., Garrido, G., & Grafman, J. (2007). Social concepts are represented in the superior anterior temporal cortex. *Proceedings of the National Academy of Sciences, 104*(15), 6430–6435.

Why do people "sacrifice life for righteousness"? An explanation from the cognitive neural mechanism of protected values

YUE Tong[1], HUANG Xiting[1], FU Anguo[2]

([1] *Research Center for Psychology and Social Development, Southwest University; School of Psychology, Southwest University, Chongqing* 400715, *China*)([2] *School of Management, Hainan University; Hainan Institute of Corporate Governance, Haikou* 570228, *China*)

Abstract: Protected values are those which people refuse to trade for any other value, especially for economic value. The evidence of cognitive neuroscience research in recent years provides two reasons why people exhibit the behavior of "sacrificing life for righteousness" for protected values, such as defying their own interests or even giving up their life. This is because protected values are characterized and constructed by an absolute rule of deontology, which makes people less likely to weigh the advantages and disadvantages of a trade; protected values are also closely related to the processes of self-identity and moral identity, and thus are subjectively endowed with the highest value. Further studies are needed to conduct empirical research on the core values of Chinese people based on the research paradigm of protective values, focus on its cultural differences, and carry out the corresponding practical research on intervention and guidance of correct values.

Key words: protected values, cognitive neural mechanism, deontological, utilitarian, backfire effect